SHANNON, CLAUDE ELWOOD (*b.* Petoskey, Michigan, 30 April 1916, *d.* Medford, Massachusetts, 24 February 2001), *engineering sciences, communication sciences, cryptography, information theory.*

Shannon is first and foremost known as a pioneer of the information age, ever since he demonstrated in his seminal paper "A Mathematical Theory of Communication" (1948) that information could be defined and measured as a scientific notion. The paper gave rise to "information theory," which includes metaphorical applications in very different disciplines, ranging from biology to linguistics via thermodynamics or quantum physics on the one hand, and a technical discipline of mathematical essence, based on crucial concepts like that of channel capacity, on the other. Shannon never showed much enthusiasm for the first kind of informal applications. He focused on the technical aspects and also contributed significantly to other fields such as cryptography, artificial intelligence, and domains where his ideas had their roots and could be readily applied in a strict fashion, that is, telecommunications and coding theory.

Formative Years. Claude Elwood Shannon was the son of Claude Shannon Sr. (1862–1934), a businessman who was also a judge of probate, and Mabel Wolf Shannon (1880–1945), a high school principal. Until the age of sixteen, he lived in Gaylord, Michigan, where his mother worked. His youth was to prove a decisive influence on his life as a scientist: his grandfather was a tinkerer, possessed a patent on a washing machine, and created various—sometimes nonsensical—objects. By the time he graduated from high school, the young Shannon had already built a radio-controlled boat and a telegraphic system to communicate with a friend nearly a mile away, using barbed wires. He made some pocket money by fixing various electrical devices, such as radios, and he admired Edison, with whom he discovered later that he shared a common ancestor.

Shannon left Gaylord in 1932 for the University of Michigan, where he studied both electrical engineering and mathematics, obtaining in 1936 a bachelor of science degree in both fields. He then found a way to match his tinkering capacities with his knowledge in electrical engineering, working in the Department of Electrical Engineering at the Massachusetts Institute of Technology (MIT) on the maintenance of the differential analyzer that had been constructed by Vannevar Bush (1890–1974). Bush was to become his mentor over the next decades. It was in Bush's department that Shannon wrote his master's thesis, titled "Symbolic Analysis of Relay and Switching Circuits," which he submitted on 10 August 1937. In an interview, Shannon recalled in 1987:

> The main machine was mechanical with spinning disks and integrators, and there was a complicated control circuit with relays. I had to understand both of these. The relay part got me interested. I knew about symbolic logic at the time from a course at Michigan, and I realized that Boolean algebra was just the thing to take care of relay circuits and switching circuits. I went to the library and got all the books I could on symbolic logic and Boolean algebra, started interplaying the two, and wrote my Master's thesis on it. That was the beginning of my great career! (Sloane and Wyner, eds., 1993, p. xxv)

The insight was decisive: It constituted "a landmark in that it helped to change digital circuit design from an art to a science" (Goldstine, 1972, p. 119). His study dealt with the circuits based on relays and switching units, such as automatic telephone exchange systems or industrial motor equipment. He developed rigorous methods for both analysis and synthesis of circuits, showing how they could be simplified. At this time, he probably had his first intuitions on the relations between redundancy and reliability, which he was to deepen later. That his stance was both theoretical and practical becomes clear at the end of his master's thesis, where he illustrated his approach with five circuits: a selective circuit, an electronic combination lock, a vote counting circuit, a base-two adder, and a factor table machine.

This dual approach was also revealed in an important letter that Shannon sent to Bush in February 1939. He wrote that "Off and on [he had] been working on an analysis of some of the fundamental properties of general systems for the transmission of intelligence, including telephony, radio, television, telegraphy, etc." He stated that "Practically all systems of communication may be thrown into the following form: $f_1(t) \rightarrow \boxed{T} \rightarrow F(t) \rightarrow \boxed{R} \rightarrow f_2(t)$ $f_1(t)$ is a general function of time (arbitrary except for certain frequency limitations) representing the intelligence to be transmitted. It represents for example, the pressure-time function in radio and telephony, or the voltage-time curve output of an iconoscope in television."

Shannon was awarded the Alfred Noble Prize of the American Society of Civil Engineers for his master's thesis in 1940. He continued to work on the use of algebra to deepen analogies and began his doctoral studies in mathematics, with the same supervisor, the algebraist Frank L. Hitchcock. The topic, however, stemmed from Bush, who suggested that Shannon apply Boolean algebra to genetics, as he had to circuits. The result of his research was submitted in the spring of 1940 in his thesis "An Algebra for Theoretical Genetics." Meanwhile, Shannon had also published his "Mathematical Theory of the Differential Analyzer" (1941) and during the summer of 1940 had started working at the Bell Laboratories, where he applied the ideas contained in his master's thesis. He also spent a few months at the Institute for Advanced Study in Princeton working under Hermann Weyl thanks to a National Research Fellowship, and he then returned to the Bell Labs, where he worked from 1941 to 1956.

The Impact of World War II. Any scientist who worked in public institutions, private companies, or universities at this time became increasingly engaged in the war effort. From 1940 onward, interdisciplinary organizations were founded: first the National Defense Research Committee (NDRC, June 1940), under the supervision of Vannevar Bush, and later the Office of Scientific Research and Development (May 1941), which included the NDRC and medical research. Shannon soon became involved in this war-related research, mainly with two projects.

The first project focused on anti-aircraft guns, which were so important in defending Great Britain under the V1 bombs and V2 rockets and more generally for air defense. Because World War II planes flew twice as high and twice as fast as those of World War I, the fire control parameters had to be automatically determined by means of radar data. Shannon was hired by Warren Weaver, at the time also head of the Natural Sciences Division of the Rockefeller Foundation. He worked with Richard B. Blackman and Hendrik Bode, also from Bell Labs. Their report, "Data Smoothing and Prediction in Fire-Control Systems," pointed in the direction of generality in signal processing. Fire control was seen as "a special case of the transmission, manipulation, and utilization of intelligence." They stated that there was "an obvious analogy between the problem of smoothing the data to eliminate or reduce the effect of tracking errors and the problem of separating a signal from interfering noise in communications systems" (Mindell, Gerovitch, and Segal, 2003, p. 73).

The second project was in the field of cryptography. At the outbreak of the war, communications could be

easily intercepted. The main transatlantic communication meant for confidential messages was the A3 telephone system developed at Bell Labs, which simply inverted parts of the bandwidth and was easily deciphered by the Germans. Shannon worked on the X-System, which solved this problem, and met British mathematician Alan Turing during this time. Turing had come to Bell Labs to coordinate British and American research on jamming, but the "need-to-know" rule that prevailed prevented them from engaging in a real exchange on these issues. The quintessence of Shannon's contribution to war cryptography can be found in a 1945 report (declassified in 1957) titled "A Mathematical Theory of Cryptography," which outlined the first theory, relying on both algebraic and probabilistic theories. Shannon explained that he was interested in discrete information consisting of sequences of discrete symbols chosen from a finite set. He gave definitions of redundancy and equivocation, and also of "information." Trying to quantify the uncertainty related to the realization of an event chosen among n events for which a probability p_i is known, he proposed the formula $H = \sum_{i=1}^{n} p_i \log p_i$ where H was at first merely a measure of uncertainty. He then showed that this formula verified eleven properties such as additivity (information brought by two selections of an outcome equals the sum of the information brought by each event) or the fact that H was maximum when all the events had the same probability (which corresponds to the worst case for deciphering). For the choice of the letter H, obviously referring to Boltzmann's H-Theorem, he explained that "most of the entropy formulas contain terms of this type" (Sloane and Wyner, 1993, pp. 84–142). According to some authors, it might have been John von Neumann who gave Shannon the following hint:

> You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one really knows what entropy really is, so in a debate you will always have the advantage. (Tribus, 1971, p. 179)

From Cryptography to Communication Theory. In his 1945 memorandum, Shannon also developed a general schema for a secured communication. The key source was represented as a disturbing element conceptualize as a "noise," similar to the message, but apart from that, the schema was similar to the one he described in 1939 in his letter to Bush. Shannon always kept this goal in mind, even when he worked in cryptology. In 1985, Shannon declared to Price "My first getting at that was information theory, and I used cryptography as a way of legitimizing the work. … For cryptography you could write up anything in any shape, which I did" (Price, 1985, p. 169).

Relying on his experience in Bell Laboratories, where he had become acquainted with the work of other telecommunication engineers such as Harry Nyquist and Ralph Hartley, Shannon published in two issues of the *Bell System Technical Journal* his paper "A Mathematical Theory of Communication." The general approach was pragmatic; he wanted to study "the savings due to statistical structure of the original message" (1948, p. 379), and for that purpose, he had to neglect the semantic aspects of information, as Hartley did for "intelligence" twenty years before (Hartley, 1928, p. 1). For Shannon, the communication process was stochastic in nature, and the great impact of his work, which accounts for the applications in other fields, was due to the schematic diagram of a general communication system that he proposed. An "information source" outputs a "message," which is encoded by a "transmitter" into the transmitted "signal." The received signal is the sum of the transmitted signal and unavoidable "noise." It is recovered as a decoded message, which is delivered to the "destination." The received signal, which is the sum between the signal and the "noise," is decoded in the "receiver" that gives the message to destination. His theory showed that choosing a good combination of transmitter and receiver makes it possible to send the message with arbitrarily high accuracy and reliability, provided the information rate does not exceed a fundamental limit, named the "channel capacity." The proof of this result was, however, nonconstructive, leaving open the problem of designing codes and decoding means that were able to approach this limit.

The paper was presented as an ensemble of twenty-three theorems that were mostly rigorously proven (but not always, hence the work of A. I. Khinchin and later A. N. Kolmogorov, who based a new probability theory on the information concept). Shannon's paper was divided into four parts, differentiating between discrete or continuous sources of information and the presence or absence of noise. In the simplest case (discrete source without noise), Shannon presented the H formula he had already defined in his mathematical theory of cryptography, which in fact can be reduced to a logarithmic mean. He defined the bit, the contraction of "binary digit" (as suggested by John W. Tukey, his colleague at Bell Labs) as the unit for information. Concepts such as "redundancy," "equivocation," or channel "capacity," which existed as common notions, were defined as scientific concepts. Shannon stated a fundamental source-coding theorem, showing that the mean length of a message has a lower limit proportional to the entropy of the source. When noise is introduced, the channel-coding theorem stated that when the entropy of the source is less than the capacity of the channel, a code exists that allows one to

transmit a message "so that the output of the source can be transmitted over the channel with an arbitrarily small frequency of errors." This programmatic part of Shannon's work explains the success and impact it had in telecommunications engineering. The turbo codes (error correction codes) achieved a low error probability at information rates close to the channel capacity, with reasonable complexity of implementation, thus providing for the first time experimental evidence of the channel capacity theorem (Berrou and Glavieux, 1996).

Another important result of the mathematical theory of communication was, in the case of a continuous source, the definition of the capacity of a channel of band W perturbed by white thermal noise power N when the average transmitter power is limited to P, given by $C = W \log \frac{P+N}{N}$ which is the formula reproduced on Shannon's gravestone. The 1948 paper rapidly became very famous; it was published one year later as a book, with a postscript by Warren Weaver regarding the semantic aspects of information.

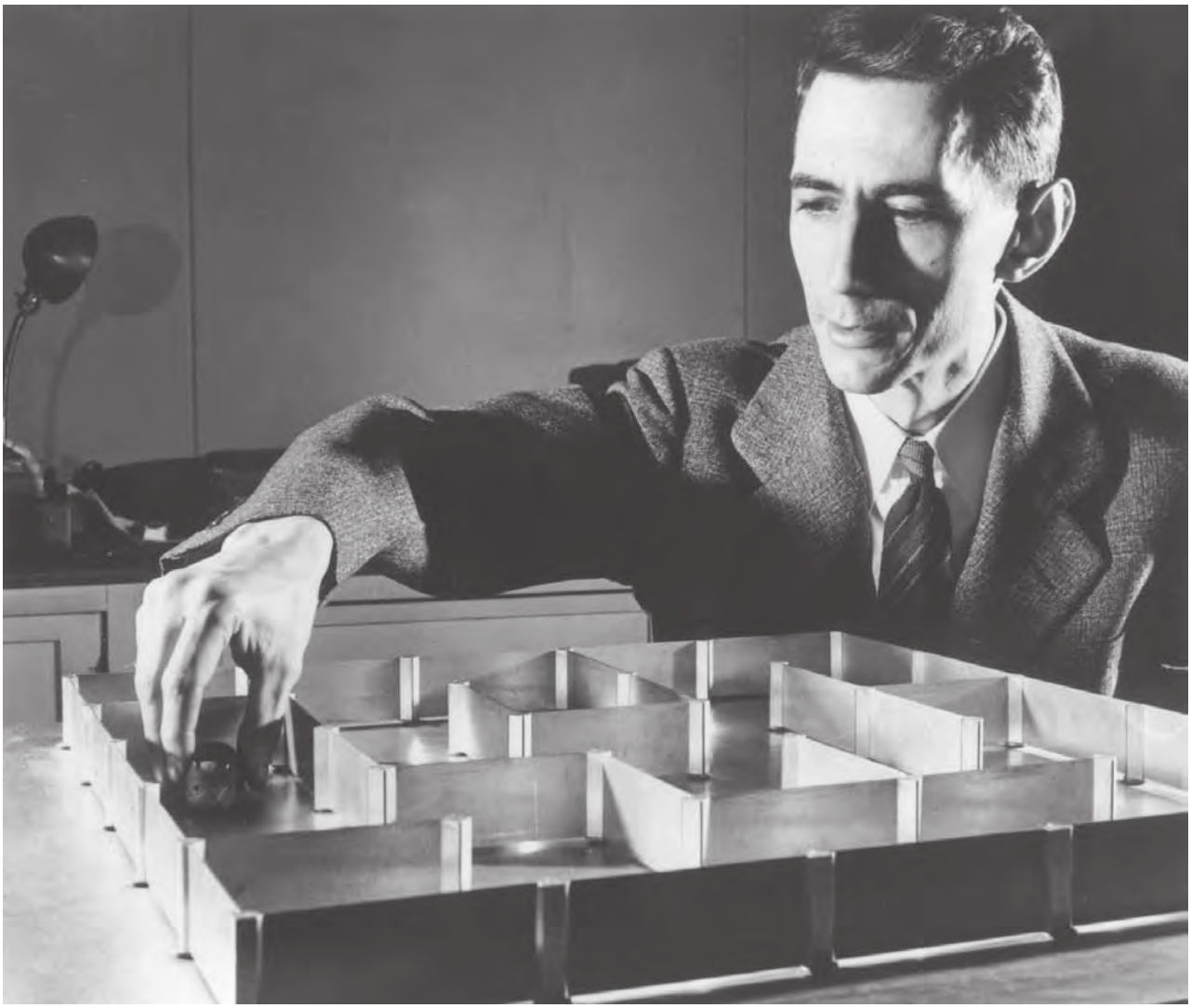
Entropy and Information. There were two different readings of this book. Some engineers became interested in the programmatic value of Shannon's writings, mostly to develop new coding techniques, whereas other scientists used the mathematical theory of communication for two reasons: on one hand, a general model of communication; and on the other, the mathematical definition of information, called "entropy" by Shannon. Those ideas coalesced with other theoretical results that appeared during the war effort, namely the idea of a general theory for "Control and Communication in the Animal and the Machine," which is the subtitle of *Cybernetics*, a book Norbert Wiener published in 1948. Shannon, von Neumann, Wiener, and others were later called "cyberneticians" during the ten meetings sponsored by the Macy Foundation, which took place between 1946 and 1953. Shannon and Weaver's 1949 book, along with the work by Wiener, brought forth a so-called "information theory."

Rapidly, connections were made between information theory and various fields, for instance in linguistics, where influences went in both directions. In order to be able to consider "natural written languages such as English, German, Chinese" as stochastic processes defined by a set of selection probabilities, Shannon relied on the work of linguists, who, in turn, were vitally interested in the calculus of the entropy of a language to gain a better understanding of concepts like that of redundancy (Shannon, 1951). Roman Jakobson was among the most enthusiastic linguists; he had participated in one of the Macy meetings in March 1948. At the very beginning of the 1950s, in most disciplines, new works were presented as "applications" of information theory, even if sometimes the application only consisted of the use of logarithmic mean. Trying to understand the connections between molecular structure and genetic information—a couple of months before the discovery of the double helix for the structure of DNA—Herman Branson calculated, in a symposium entitled "The Use of Information Theory in Biology," the information quantity (H) contained in a human. He gave the expression "H(food and environment) = H(biological function) + H(maintenance and repair) + H(growth, differentiation, memory)" (Quastler, 1953, p. 39). Henry Quastler came to the conclusion, as did Sidney Dancoff, that "H(man)" was about 2×10^{28} bits (p. 167).

Taking issue with these different kinds of applications, Shannon in 1956 wrote a famous editorial, published in the *Transactions of the Institute of Radio Engineers*, with the title "The Bandwagon." As he stated, referring to his 1948 paper, "Starting as a technical tool for the communication engineer, it has received an extraordinary amount of publicity in the popular as well as the scientific press. In part, this has been due to connections with such fashionable fields as computing machines, cybernetics, and automation; and in part, to the novelty of its subject matter. As a consequence, it has perhaps been ballooned to an importance beyond its actual accomplishments." At this time, some applications of information theory already reflected a mood, essentially based on a loose, rather than a scientific definition of information. Forty years later, the project of "information highways," presented to promote the Internet, partly relied on the same idea.

Shannon as a Pioneer in Artificial Intelligence. At the time Shannon published his relatively pessimistic editorial, he was already engaged in other research, typically related to his ability to combine mathematical theories, electrical engineering, and "tinkering," namely, artificial intelligence. Shannon coauthored the 1955 "Proposal for the Dartmouth Summer Research Project on Artificial Intelligence," which marked the debut of the term "artificial intelligence." Together with Nathaniel Rochester, John McCarthy, and Marvin L. Minsky, he obtained support from the Rockefeller Foundation to "proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it." In explaining his own goal, Shannon named two topics.

The first topic, presented as an "application of information theory," was based on an analogy: in the same way that information theory was concerned with the reliable transmission of information over a noisy channel, he wanted to tackle the structure of computing machines in which reliable computing is supposed to be achieved using some unreliable elements, a problem to which John von Neumann devoted considerable attention. Starting from



Claude Shannon. *Claude Shannon with an electronic mouse which has a "super" memory and can learn its way round a maze without a mistake after only one "training" run.* **HULTON ARCHIVE/GETTY IMAGES.**

this parallel, notions such as redundancy and channel capacity were to be used to improve the architecture of computing machines.

The second topic dealt with the way in which a "brain model" can adapt to its environment. This had no direct link with information theory but was more related to the work Shannon had presented during the eighth Macy meeting, in March 1951, where he gathered with other cyberneticians. Shannon demonstrated an electromechanical mouse he called Theseus, which would be "taught" to find its way in a labyrinth. In his Dartmouth proposal, Shannon put the emphasis on "clarifying the environmental model, and representing it as a mathematical structure." He had already noticed that "in discussing mechanized intelligence, we think of machines performing the most advanced human thought activities—proving theorems, writing music, or playing chess." He posited a bottom-up approach in the "direction of these advanced activities," starting with simpler models, as he had done in his 1950 paper entitled "Programming a Computer for Playing Chess." In this first published article on computer chess, Shannon offered the key elements for writing a "program," such as an "evaluation function" or a "minimax procedure."

A Complex Legacy. Shannon's contributions to artificial intelligence have often been neglected because of the enormous aura. He is so well known for his work on information theory that his credit for AI is often ignored. Most history of AI does not even mention his presence at the

Dartmouth meeting of information theory. None of the works he wrote after the 1950s received such recognition. He left Bell Labs for the Massachusetts Institute of Technology (MIT) in 1956, first as a visiting professor; he was a permanent member of the Research Laboratory of Electronics at MIT for twenty years, starting in 1958, after he had spent a year as a fellow at the Center for Advanced Study in the Behavioral Sciences in Palo Alto.

Most of his scientific work was devoted to the promotion and deepening of information theory. Shannon was invited to many countries, including the Soviet Union in 1965. While there, giving a lecture at an engineering conference, he had an opportunity to play a chess match against Mikhail Botvinik. He tackled the case of transmission with a memoryless channel (a noisy channel where the noise acts independently on each symbol transmitted through the channel). It is on this topic that he published his last paper related to information theory, as early as 1967, with Robert G. Gallager and Elwyn R. Berlekamp.

In the late 1960s and 1970s, Shannon became interested in portfolio management and, more generally, investment theory. One of his colleagues at Bell Labs, John L. Kelly, had shown in 1956 how information theory could be applied to gambling. Together with Ed Thorp, Shannon went to Las Vegas to test their ideas. In 1966 they also invented the first wearable computer at MIT that was able to predict roulette wheels.

Shannon never gave up constructing eccentric machines, like the THROBAC (THrifty ROman-numeral BAckward-looking Computer) he built in the 1950s, the rocket-powered Frisbee, or a device that could solve the Rubik's Cube puzzle. He developed many automata, many of which he kept at his home: among others, a tiny stage on which three clowns could juggle with eleven rings, seven balls, and five clubs, all driven by an invisible mechanism of clockwork and rods. Juggling was one of his passions, which also included playing chess, riding a unicycle, and playing to clarinet. In the early 1980s Shannon began writing an article for *Scientific American* called "Scientific Aspects of Juggling," which he never finished (Sloane and Wyner, 1993, pp. 850–864).

At the dawn of the twenty-first century, Shannon's contributions are manifold. Whereas there are still applications that only consist of using the logarithmic mean or the schematic diagram of a general communication system (applications he condemned in his 1956 editorial, "The Bandwagon"), there are also numerous new fields that could not be defined without referring to his work. In the field of technology, coding theories that are applied to compact discs or deep-space communication are merely developments of information theory. In mathematics, entire parts of algorithmic complexity theory can be seen as resulting from the development of Shannon's theory. In biology, the protean use made of the expression "genetic information" explains the development of molecular biology (Fox Keller, Kay and Yockey). From the 1990s onward, in physics, the domain of "quantum information" took off around the definition of *qubits,* which extended the *bit* initially used by Shannon to measure information. Shannon unfortunately could not take part in these developments nor take them into account; from the mid-1990s he struggled with Alzheimer's disease, to which he succumbed in February 2001.

BIBLIOGRAPHY

A comprehensive bibliography appears in Neil J. A. Sloane and Aaaron D. Wyner, eds., Claude Elwood Shannon: Collected Papers, *Piscataway, NJ: IEEE Press, 1993. These collected papers include the 1937 master's thesis (http://libraries.mit.edu/); the "Letter to Vannevar Bush, Feb. 16, 1939"; and the 1940 PhD dissertation (http://libraries.mit.edu/). The master's and PhD essays are also available at the MIT's online institutional repository http://dspace.mit.edu/handle/1721.1/11173 and http://dspace.mit.edu/handle/1721.1/11174;the 1939 letter was first reproduced in Hagmeyer's doctoral dissertation (see below). Shannon's archives are at the Bell Laboratories Archives and at the National Archives in Washington, DC*

WORKS BY SHANNON

"A Mathematical Theory of Communication." *Bell System Technical Journal* 27 (1948): 379–423, 623–656.

"Communication in the Presence of Noise." *Proceedings of the Institute of Radio Engineers* 37 (1949): 10–21.

"Communication Theory of Secrecy Systems." *Bell System Technical Journal* 28 (1949): 656–715.

With Warren Weaver. *The Mathematical Theory of Communication.* Urbana: University of Illinois Press, 1949.

"Programming a Computer for Playing Chess." *Philosophical Magazine* 41 (1950): 256–275.

"Prediction and Entropy in Printed English." *Bell System Technical Journal* 30 (1951): 50–64.

"The Bandwagon." *IRE Transactions on Information Theory* 2 (1956): 3.

With Robert G. Gallager and Elwyn R. Berlekamp. "Lower Bounds to Error Probability for Coding on Discrete Memorylless Channels." *Information and Control* 10 (1967): 65-103.

OTHER SOURCES

Berrou, Claude, and Alain Glavieux. "Near Optimum Error Correcting Coding and Decoding: Turbo-Codes." *IEEE Transactions on Communications* 44 (1996): 1261–1271.

Foerster, Heinz von. *Cybernetics, Circular Causal and Feedback Mechanisms in Biological and Social Systems.* New York: Macy Foundation, 1952.

Fox Keller, Evelyn. *The Century of the Gene.* Cambridge, MA: Harvard University Press, 2000.

Goldstine, Herman H. *The Computer from Pascal to von Neumann.* Princeton, NJ: Princeton University Press, 1972.

Hagemeyer, Friedrich W. *Die Entstehung von Informationskonzepten in der Nachrichtentechnik*. Doktorarbeit an der Freie Universität Berlin (PhD), FB 11 Philosophie und Sozialwissenschaften, 1979.

Hartley, Ralph V. L. "Transmission of Information." *Bell System Technical Journal* 7 (1928): 535–563.

Hodges, Andrew. *Alan Turing: The Enigma*. London: Burnett Books, 1983.

Horgan, John. "Claude E. Shannon: Unicyclist, Juggler and Father of Information Theory." *Scientific American* 242 (1990): 20–22B.

Kay, Lily E. *Who Wrote the Book of Life? A History of the Genetic Code*. Chicago: University of Chicago Press, 2000.

Kelly, John L. "A New Interpretation of the Information Rate." *Bell System Technical Journal* 35 (1956): 917–925.

Mindell, D., S. Gerovitch, and J. Segal. "From Communications Engineering to Communications Science: Cybernetics and Information Theory in the United States, France, and the Soviet Union." In *Science and Ideology: A Comparative History*, edited by Mark Walker, pp. 66–96. London: Routledge, 2003.

Pias, Claus. *Cybernetics/Kybernetik. The Macy-Conferences, 1946–1953. Transactions/Protokolle*. 2 vol. Fernwald, Germany: Diaphanes, 2003 and 2004.

Price, Robert. "A Conversation with Claude Shannon: One Man's Approach to Problem Solving." *Cryptologia* 9 (1985): 167–175.

Quastler, Henry, ed. *Essays on the Use of Information Theory in Biology*. Urbana: University of Illinois Press, 1953.

Segal, Jérôme. *Le Zéro et le un: Histoire de la notion scientifique d'information*. Paris: Syllepse, 2003.

Tribus, Myron, and E. C. McIrvine. "Energy and Information." *Scientific American* 224 (1971): 178–184.

Verdu, Sergio. "Fifty Years of Shannon Theory." *IEEE Transactions on Information Theory* 44 (1998): 2057–2078.

Wiener, Norbert. *Cybernetics , or Control and Communication in the Animal and the Machine*. Paris, France: Hermann et Cie, 1948.

Yockey, Hubert P. *Information Theory and Molecular Biology*. Cambridge, U.K.: Cambridge University Press, 1992.

———. *Information Theory, Evolution, and the Origin of Life*. Cambridge, U.K.: Cambridge University Press, 2005.

Jérôme Segal